* Potential topics:
  + Subject matter interest by congressman
  + Power of congressman
  + Number of bills about congressional rules
  + Prevalence of topic by year
  + Interactive of prevalence of bill topic
* Title: Congressional BEAT (Bills Exploratory Analysis and Tracking)
* Background and Motivation
  + We want to create resources to allow the public to better understand how congress works. There are an overwhelming number of bills introduced each year (>6,000) and many do not know what they contain and who sponsors them. Given the public distaste for the way media presents politics, it is becoming more important for individuals to understand and be given raw information to draw their own conclusions. Normally the media is focused on one or two bills, despite many being discussed concurrently.
* Project Objectives
  + What are the scientific and inferential goals for this project?
  + What would you like to learn and accomplish?
    - Prevalence of topic by year;
    - prevalence of bill topic in general; likelihood of a certain bill topic to be voted on; likelihood of a certain topic to become a law
    - Geographic prevalence of bills by state – in title alone; Geographic display of number of bills introduced by each state
    - Subject matter interest by congressman
  + List the benefits.: This is a very opaque thing to most people, even if you’re interested in it. This allows the public to more clearly understand what’s happening overall. There’s a lot of detail to sift through if just looking at bills individually.
  + Optional features:
    - Legislators with similar interests – Look at the how many bills of each type of subject are sponsored by the same congressman. For example, a constituent is concerned about child welfare; which congressman is most likely to be interested in that topic?
    - Conduct a network analysis of bill co-sponsorship
    - Dig deeper into resolutions and how they influence the movement of bills (especially by coupling unrelated bills into the same resolution)
* What Data
  + We are using congressional records from Congress.gov. They provide downloadable, bulk data with information on bills, timelines and voting records. <https://github.com/usgpo/bill-status/blob/master/BILLSTATUS-XML_User_User-Guide.md>. We may also utilize voting records (<http://clerk.house.gov/legislative/legvotes.aspx>)
* Design Overview
  + List the statistical and computational methods you plan to use.
    - Scraping data from the internet
    - Exploratory Data Analysis
    - Logistics/Poisson Regression
    - Classification and Clustering
    - Data Visualization (eg. Heap map, word cloud)
* Schedule/timeline
  + Weekly meetings at 9:30 am on Tuesdays, and additional meetings as needed.
  + Assigned modules:
    - Hillary: Geographic prevalence of bills by state – in title alone; geographic visualization regarding how many bills are sponsored from each state; Screen cast video
    - Liam: Scraping the data; Subject matter interest by congressman
    - Sara: Prevalence of topic by year; prevalence of bill topic in general; likelihood of a certain bill topic to be voted on; likelihood of a certain topic to become a law
    - Website: all
  + By 11/5:
    - Set up a Github
    - scrape the data
  + By 11/12:
    - Learn how to use R-shiny
    - Write coding to get summary statistics
    - Exploration of the data
  + By 11/19:
    - Complete analysis work for each module
  + By 11/26:
    - Complete data visualizations and begin working on write up
  + By 12/3:
    - Complete write up to accompany data visualizations, and incorporate them to the website
  + By 12/5:
* Make the Screen Cast video

Due on 12/8